

CW

ENTERED

see page 6&7



PCT10

RAW SEQUENCE LISTING

DATE: 01/21/2003

PATENT APPLICATION: US/10/019,258

TIME: 13:57:52

Input Set : A:\EP.txt

Output Set: N:\CRF4\01212003\J019258.raw

3 <110> APPLICANT: Lexow, Preben
 5 <120> TITLE OF INVENTION: Method of cloning and producing fragment chains with
 readable information

6 content
 8 <130> FILE REFERENCE: 1181-256
 10 <140> CURRENT APPLICATION NUMBER: US 10/019258
 11 <141> CURRENT FILING DATE: 2001-12-28
 13 <150> PRIOR APPLICATION NUMBER: PCT/GB00/02512
 14 <151> PRIOR FILING DATE: 2000-06-27
 16 <150> PRIOR APPLICATION NUMBER: NO 20003191
 17 <151> PRIOR FILING DATE: 2000-06-20
 19 <150> PRIOR APPLICATION NUMBER: NO 20003190
 20 <151> PRIOR FILING DATE: 2000-06-20
 22 <150> PRIOR APPLICATION NUMBER: NO 19991325
 23 <151> PRIOR FILING DATE: 1999-06-28
 25 <160> NUMBER OF SEQ ID NOS: 105
 27 <170> SOFTWARE: PatentIn version 3.1
 29 <210> SEQ ID NO: 1
 30 <211> LENGTH: 11
 31 <212> TYPE: DNA
 32 <213> ORGANISM: Artificial Sequence
 34 <220> FEATURE:
 35 <223> OTHER INFORMATION: Adapter
 37 <220> FEATURE:
 38 <221> NAME/KEY: misc_feature
 39 <222> LOCATION: (8)..(9)
 40 <223> OTHER INFORMATION: N is any nucleotide.
 43 <400> SEQUENCE: 1

W--> 44 ggcccccnna a

11

47 <210> SEQ ID NO: 2
 48 <211> LENGTH: 11
 49 <212> TYPE: DNA
 50 <213> ORGANISM: Artificial Sequence
 52 <220> FEATURE:
 53 <223> OTHER INFORMATION: Adapter
 55 <220> FEATURE:
 56 <221> NAME/KEY: misc_feature
 57 <222> LOCATION: (7)..(9)
 58 <223> OTHER INFORMATION: N is any nucleotide.
 61 <400> SEQUENCE: 2

W--> 62 ggggcnnc t

11

65 <210> SEQ ID NO: 3
 66 <211> LENGTH: 23
 67 <212> TYPE: DNA

RAW SEQUENCE LISTING

DATE: 01/21/2003

PATENT APPLICATION: US/10/019,258

TIME: 13:57:52

Input Set : A:\EP.txt

Output Set: N:\CRF4\01212003\J019258.raw

```

68 <213> ORGANISM: Artificial Sequence
70 <220> FEATURE:
71 <223> OTHER INFORMATION: BbvI overhang
73 <400> SEQUENCE: 3
74 cgagcgcctc cagtgcagcg gag 23
77 <210> SEQ ID NO: 4
78 <211> LENGTH: 24
79 <212> TYPE: DNA
80 <213> ORGANISM: Artificial Sequence
82 <220> FEATURE:
83 <223> OTHER INFORMATION: BbvI overhang
85 <400> SEQUENCE: 4
86 tatcgcgcct ccagtgcagc ggag 24
89 <210> SEQ ID NO: 5
90 <211> LENGTH: 24
91 <212> TYPE: DNA
92 <213> ORGANISM: Artificial Sequence
94 <220> FEATURE:
95 <223> OTHER INFORMATION: BbvI overhang
97 <400> SEQUENCE: 5
98 ctctgcgcct ccagtgcagc ggag 24
101 <210> SEQ ID NO: 6
102 <211> LENGTH: 24
103 <212> TYPE: DNA
104 <213> ORGANISM: Artificial Sequence
106 <220> FEATURE:
107 <223> OTHER INFORMATION: BbvI overhang 6 (delC)
109 <400> SEQUENCE: 6
110 ctctctccgc tgcactggag gcgc 24
113 <210> SEQ ID NO: 7
114 <211> LENGTH: 24
115 <212> TYPE: DNA
116 <213> ORGANISM: Artificial Sequence
118 <220> FEATURE:
119 <223> OTHER INFORMATION: BbvI overhang 7a
121 <400> SEQUENCE: 7
122 caacgcgcct ccagtgcagc ggag 24
125 <210> SEQ ID NO: 8
126 <211> LENGTH: 24
127 <212> TYPE: DNA
128 <213> ORGANISM: Artificial Sequence
130 <220> FEATURE:
131 <223> OTHER INFORMATION: BbvI overhang 9b
133 <400> SEQUENCE: 8
134 ggtagcgcct ccagtgcagc ggag 24
137 <210> SEQ ID NO: 9
138 <211> LENGTH: 25
139 <212> TYPE: DNA
140 <213> ORGANISM: Artificial Sequence

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/019,258

DATE: 01/21/2003

TIME: 13:57:52

Input Set : A:\EP.txt

Output Set: N:\CRF4\01212003\J019258.raw

```

142 <220> FEATURE:
143 <223> OTHER INFORMATION: Cloning site 1a
145 <400> SEQUENCE: 9
146 aagagctccg ctgcactgga ggcgc 25
149 <210> SEQ ID NO: 10
150 <211> LENGTH: 25
151 <212> TYPE: DNA
152 <213> ORGANISM: Artificial Sequence
154 <220> FEATURE:
155 <223> OTHER INFORMATION: Cloning site 1b
157 <400> SEQUENCE: 10
158 ctcttctccg ctgcactgga ggcgc 25
161 <210> SEQ ID NO: 11
162 <211> LENGTH: 35
163 <212> TYPE: DNA
164 <213> ORGANISM: Artificial Sequence
166 <220> FEATURE:
167 <223> OTHER INFORMATION: Consensus binding motifs of the initiation linkers
169 <220> FEATURE:
170 <221> NAME/KEY: misc_feature
171 <222> LOCATION: (19)..(24)
172 <223> OTHER INFORMATION: N is any nucleotide.
175 <400> SEQUENCE: 11
W--> 176 gcagcgacca tgagtccanc tcnngtggat gacgc 35
179 <210> SEQ ID NO: 12
180 <211> LENGTH: 37
181 <212> TYPE: DNA
182 <213> ORGANISM: Artificial Sequence
184 <220> FEATURE:
185 <223> OTHER INFORMATION: Initiation linker
187 <220> FEATURE:
188 <221> NAME/KEY: misc_feature
189 <222> LOCATION: (19)..(37)
190 <223> OTHER INFORMATION: N is any nucleotide with the proviso that the DNA sequence
from 3
191 2 to 37 is not palindromic.
194 <400> SEQUENCE: 12
W--> 195 gcagcgacca tgagtccanc tcnngtggat gnnnnnn 37
198 <210> SEQ ID NO: 13
199 <211> LENGTH: 38
200 <212> TYPE: DNA
201 <213> ORGANISM: Artificial Sequence
203 <220> FEATURE:
204 <223> OTHER INFORMATION: Initiation linker
206 <220> FEATURE:
207 <221> NAME/KEY: misc_feature
208 <222> LOCATION: (19)..(38)
209 <223> OTHER INFORMATION: N is any nucleotide with the proviso that the DNA sequence
from 3
210 3 to 38 is not palindromic.
213 <400> SEQUENCE: 13

```

RAW SEQUENCE LISTING

DATE: 01/21/2003

PATENT APPLICATION: US/10/019,258

TIME: 13:57:52

Input Set : A:\EP.txt

Output Set: N:\CRF4\01212003\J019258.raw

```

W--> 214 gcagcgacca tgagtccanc tcnngtggat gnnnnnnnn      38
217 <210> SEQ ID NO: 14
218 <211> LENGTH: 39
219 <212> TYPE: DNA
220 <213> ORGANISM: Artificial Sequence
222 <220> FEATURE:
223 <223> OTHER INFORMATION: Initiation linker
225 <220> FEATURE:
226 <221> NAME/KEY: misc_feature
227 <222> LOCATION: (19)..(39)
228 <223> OTHER INFORMATION: N is any nucleotide with the proviso that the DNA sequence
from 3
229      4 to 39 is not palindromic.
232 <400> SEQUENCE: 14
W--> 233 gcagcgacca tgagtccanc tcnngtggat gnnnnnnnnnn      39
236 <210> SEQ ID NO: 15
237 <211> LENGTH: 40
238 <212> TYPE: DNA
239 <213> ORGANISM: Artificial Sequence
241 <220> FEATURE:
242 <223> OTHER INFORMATION: Initiation linker
244 <220> FEATURE:
245 <221> NAME/KEY: misc_feature
246 <222> LOCATION: (19)..(40)
247 <223> OTHER INFORMATION: N is any nucleotide with the proviso that the DNA sequence
from 3
248      5 to 40 is not palindromic.
251 <400> SEQUENCE: 15
W--> 252 gcagcgacca tgagtccanc tcnngtggat gnnnnnnnnnn      40
255 <210> SEQ ID NO: 16
256 <211> LENGTH: 41
257 <212> TYPE: DNA
258 <213> ORGANISM: Artificial Sequence
260 <220> FEATURE:
261 <223> OTHER INFORMATION: Initiation linker
263 <220> FEATURE:
264 <221> NAME/KEY: misc_feature
265 <222> LOCATION: (19)..(41)
266 <223> OTHER INFORMATION: N is any nucleotide with the proviso that the DNA sequence
from 3
267      6 to 41 is not palindromic.
270 <400> SEQUENCE: 16
W--> 271 gcagcgacca tgagtccanc tcnngtggat gacgcnnnnnn n      41
274 <210> SEQ ID NO: 17
275 <211> LENGTH: 42
276 <212> TYPE: DNA
277 <213> ORGANISM: Artificial Sequence
279 <220> FEATURE:
280 <223> OTHER INFORMATION: Initiation linker
282 <220> FEATURE:
283 <221> NAME/KEY: misc_feature
284 <222> LOCATION: (19)..(42)

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/019,258

DATE: 01/21/2003

TIME: 13:57:52

Input Set : A:\EP.txt

Output Set: N:\CRF4\01212003\J019258.raw

```

      285 <223> OTHER INFORMATION: N is any nucleotide with the proviso that the DNA sequence
from 3
      286           7 to 42 is not palindromic.
      289 <400> SEQUENCE: 17
W--> 290 gcagcgacca tgagtccanc tcnngtggat gacgcnnnnn nn                42
      293 <210> SEQ ID NO: 18
      294 <211> LENGTH: 43
      295 <212> TYPE: DNA
      296 <213> ORGANISM: Artificial Sequence
      298 <220> FEATURE:
      299 <223> OTHER INFORMATION: Initiation linker
      301 <220> FEATURE:
      302 <221> NAME/KEY: misc_feature
      303 <222> LOCATION: (19)..(43)
      304 <223> OTHER INFORMATION: N is any nucleotide with the proviso that the DNA sequence
from 3
      305           8 to 43 is not palindromic.
      308 <400> SEQUENCE: 18
W--> 309 gcagcgacca tgagtccanc tcnngtggat gacgcnnnnn nnn            43
      312 <210> SEQ ID NO: 19
      313 <211> LENGTH: 44
      314 <212> TYPE: DNA
      315 <213> ORGANISM: Artificial Sequence
      317 <220> FEATURE:
      318 <223> OTHER INFORMATION: Initiation linker
      320 <220> FEATURE:
      321 <221> NAME/KEY: misc_feature
      322 <222> LOCATION: (19)..(44)
      323 <223> OTHER INFORMATION: N is any nucleotide with the proviso that the DNA sequence
from 3
      324           9 to 44 is not palindromic.
      327 <400> SEQUENCE: 19
W--> 328 gcagcgacca tgagtccanc tcnngtggat gacgcnnnnn nnnn          44
      331 <210> SEQ ID NO: 20
      332 <211> LENGTH: 45
      333 <212> TYPE: DNA
      334 <213> ORGANISM: Artificial Sequence
      336 <220> FEATURE:
      337 <223> OTHER INFORMATION: Initiation linker
      339 <220> FEATURE:
      340 <221> NAME/KEY: misc_feature
      341 <222> LOCATION: (19)..(45)
      342 <223> OTHER INFORMATION: N is any nucleotide with the proviso that the DNA sequence
from 4
      343           0 to 45 is not palindromic.
      346 <400> SEQUENCE: 20
W--> 347 gcagcgacca tgagtccanc tcnngtggat gacgcnnnnn nnnnn         45
      350 <210> SEQ ID NO: 21
      351 <211> LENGTH: 46
      352 <212> TYPE: DNA
      353 <213> ORGANISM: Artificial Sequence
      355 <220> FEATURE:
      356 <223> OTHER INFORMATION: Initiation linker

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/019,258

DATE: 01/21/2003
TIME: 13:57:53

Input Set : A:\EP.txt
Output Set: N:\CRF4\01212003\J019258.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:1; N Pos. 8,9
Seq#:2; N Pos. 7,8,9
Seq#:11; N Pos. 19,23,24
Seq#:12; N Pos. 19,23,24,32,33,34,35,36,37
Seq#:13; N Pos. 19,23,24,32,33,34,35,36,37,38
Seq#:14; N Pos. 19,23,24,32,33,34,35,36,37,38,39
Seq#:15; N Pos. 19,23,24,32,33,34,35,36,37,38,39,40
Seq#:16; N Pos. 19,23,24,36,37,38,39,40,41
Seq#:17; N Pos. 19,23,24,36,37,38,39,40,41,42
Seq#:18; N Pos. 19,23,24,36,37,38,39,40,41,42,43
Seq#:19; N Pos. 19,23,24,36,37,38,39,40,41,42,43,44
Seq#:20; N Pos. 19,23,24,36,37,38,39,40,41,42,43,44,45
Seq#:21; N Pos. 19,23,24,36,37,38,39,40,41,42,43,44,45,46
Seq#:43; N Pos. 20,21,22,23,24
Seq#:69; N Pos. 8,9,13
Seq#:70; N Pos. 1,9,10,14
Seq#:71; N Pos. 1,2,10,11,15
Seq#:72; N Pos. 1,2,3,11,12,16
Seq#:73; N Pos. 12,13,17
Seq#:74; N Pos. 1,13,14,18
Seq#:75; N Pos. 1,2,14,15,19
Seq#:76; N Pos. 1,2,3,15,16,20
Seq#:77; N Pos. 1,2,3,4,16,17,21
Seq#:78; N Pos. 1,2,3,4,5,17,18,22
Seq#:79; N Pos. 1,2,3,4,5
Seq#:82; N Pos. 4,5,6,7,8,9
Seq#:83; N Pos. 4,5,6,7,8,9,10
Seq#:84; N Pos. 4,5,6,7,8,9
Seq#:85; N Pos. 1,2,3,4,5,6,7,8,9,17,18,22
Seq#:86; N Pos. 1,2,3,4,5,6,7,8,13,20,21,22,23,24,25,32,33,38,39,40,41,42
Seq#:86; N Pos. 43,44,45,46,47
Seq#:87; N Pos. 1,2,3,4,5,6,7,8,9,10,15,16,23,24,25,26,27,28,35,40,41,42,43
Seq#:87; N Pos. 44,45,46,47
Seq#:88; N Pos. 5,12,13,14,15,16,17,24,25
Seq#:89; N Pos. 5,6,13,14,15,16,17,18,25
Seq#:90; N Pos. 1,2,3,4,5,6,7,8,13,14,15,16,17,18,19,20,21,22
Seq#:91; N Pos. 1,2,3,4,5,6,7,8,9,10,15,16,17,18,19,20,21,22
Seq#:92; N Pos. 1,2,3,4,5,6,7,8,12,13,14,15,16,17,18,19,20,23,24,25,26,27
Seq#:92; N Pos. 32,33,34,35,36,37,38,42,43,44,45,46,47,48,49,50,51
Seq#:93; N Pos. 1,2,3,4,5,6,7,8,9,10,14,15,16,17,18,19,20,25,26,27,28,29,32
Seq#:93; N Pos. 33,34,35,36,37,38,39,40,44,45,46,47,48,49,50,51
Seq#:94; N Pos. 1,2,3,4,5,6,7,8,9,12,13,14,15,16,21,22,23,24,25,26,27
Seq#:95; N Pos. 1,2,3,4,5,6,7,12,13,14,15,16,19,20,21,22,23,24,25,26,27
Seq#:96; N Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 01/21/2003
PATENT APPLICATION: US/10/019,258 TIME: 13:57:53

Input Set : A:\EP.txt
Output Set: N:\CRF4\01212003\J019258.raw

Seq#:96; N Pos. 24,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44
Seq#:97; N Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,22,23,24,25,26

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:1; Line(s) 5

VERIFICATION SUMMARY

DATE: 01/21/2003

PATENT APPLICATION: US/10/019,258

TIME: 13:57:53

Input Set : A:\EP.txt

Output Set: N:\CRF4\01212003\J019258.raw

L:44 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:0
L:62 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:0
L:176 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11 after pos.:0
L:195 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12 after pos.:0
L:214 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13 after pos.:0
L:233 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14 after pos.:0
L:252 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15 after pos.:0
L:271 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16 after pos.:0
L:290 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17 after pos.:0
L:309 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:18 after pos.:0
L:328 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:19 after pos.:0
L:347 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20 after pos.:0
L:366 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21 after pos.:0
L:636 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:43 after pos.:0
L:954 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:69 after pos.:0
L:972 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:70 after pos.:0
L:990 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:71 after pos.:0
L:1008 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:72 after pos.:0
L:1026 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:73 after pos.:0
L:1044 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:74 after pos.:0
L:1062 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:75 after pos.:0
L:1080 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:76 after pos.:0
L:1098 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:77 after pos.:0
L:1116 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:78 after pos.:0
L:1134 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:79 after pos.:0
L:1177 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:82 after pos.:0
L:1196 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:83 after pos.:0
L:1215 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:84 after pos.:0
L:1235 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:85 after pos.:0
L:1254 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:86 after pos.:0
L:1273 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:87 after pos.:0
L:1292 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:88 after pos.:0
L:1311 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:89 after pos.:0
L:1330 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:90 after pos.:0
L:1349 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:91 after pos.:0
L:1368 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:92 after pos.:0
L:1387 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:93 after pos.:0
L:1405 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:94 after pos.:0
L:1423 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:95 after pos.:0
L:1441 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:96 after pos.:0
L:1459 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:97 after pos.:0